



Strong convergence of a class of non-homogeneous Markov arrival processes to a Poisson process

James Ledoux

► To cite this version:

James Ledoux. Strong convergence of a class of non-homogeneous Markov arrival processes to a Poisson process. *Statistics and Probability Letters*, 2008, 78 (4), pp.445-455. 10.1016/j.spl.2007.07.018 . hal-00359796

HAL Id: hal-00359796

<https://hal.science/hal-00359796>

Submitted on 26 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Strong convergence of a class of non-homogeneous Markov Arrivals Processes to a Poisson process

James LEDOUX*

22 December 2006

Abstract

In this paper, we are concerned with a time-inhomogeneous version of the Markovian Arrival Process. Under the assumption that the environment process is asymptotically time-homogeneous, we discuss a Poisson approximation of the counting process of arrivals when the arrivals are rare. We provide a rate of convergence for the distance in variation. Poisson-type approximation for the process resulting of a special marking procedure of the arrivals is outlined.

Keywords: Markov Additive Process, Compound Poisson approximation.

1 Introduction

The so-called Markovian Arrival Process (MAP) is a very popular model of arrivals process in queueing theory (see e.g. [13]). Such a process is an instance of *Markov additive process* [14]. It may be described as a bivariate random process $\{(N_t, X_t)\}_{t \geq 0}$ where: $\{N_t\}_{t \geq 0}$ is the counting process of arrivals and $\{X_t\}_{t \geq 0}$ is a finite Markov process sometimes called the “environment process”. With Lucantoni’s formalism, a MAP is parametrized by two matrices D_0 and D_1 where the non-negative entry $D_1(i, j) \geq 0$ (resp. $D_0(i, j) \geq 0, i \neq j$) represents for the environment process, the infinitesimal intensity of a jump from state e_i to e_j with one (resp. no) arrival. Moreover, $D_0 + D_1$ is the intensity matrix of the (homogeneous) Markov process $\{X_t\}_{t \geq 0}$. In this paper, we consider a class of MAP for which D_0 and D_1 are time-dependent, so that $\{(N_t, X_t)\}_{t \geq 0}$ and $\{X_t\}_{t \geq 0}$ are non-homogeneous Markov processes. A formal definition is given in Subsection 2.1. For such a class of non-homogeneous Markovian Arrival Process, we are concerned with the following question: what is the asymptotic distribution of $\{N_t\}_{t \geq 0}$ when the arrivals tend to be rare given that the environment process tend to be stationary as time elapses. A simple way is to replace the matrices $D_0(t), D_1(t)$ by perturbed matrices $D_0^{(\varepsilon)}(t), D_1^{(\varepsilon)}(t)$ which are supposed to be of the following form:

$$D_1^{(\varepsilon)}(t) = \varepsilon B_1(t) \quad \text{and} \quad D_0^{(\varepsilon)}(t) = Q(t) + \varepsilon B_0(t) \quad (1)$$

*IRMAR UMR-CNRS 6625 & INSA 20 avenue des Buttes de Coësmes, 35708 Rennes Cedex 7, FRANCE; email: james.ledoux@insa-rennes.fr

so that $\{Q(t)\}_{t \geq 0}$ and $\{D_0^{(\varepsilon)}(t) + D_1^{(\varepsilon)}(t)\}_{t \geq 0}$ are the intensity matrices of finite non-homogeneous Markov processes and $Q(t)$ converges to a stationary irreducible intensity matrix Q . Then, we investigate the asymptotic distribution of the counting process of arrivals when ε tends to 0. In Subsection 2.2, we state that the convergence in variation to a homogeneous Poisson process takes place at the specific time scale t/ε . A rate of convergence is provided. The main assumption is on the rate of the convergence of $Q(t)$ to Q . A brief discussion on such an issue is reported in Appendix A. A simple extension to the case of marked arrivals is considered in Section 3. The main result is proved in Section 4. The present work may be thought of as the natural continuation of that reported in [7]. Indeed, the present issue was discussed in [7] for specific reliability models for which the perturbed model looks like the model in (1) (in fact, it was quadratic in ε). Unlike to the present paper, these specific MAPs were assumed to be time-homogeneous.

Basic notation

By convention, vectors are row vectors. The column vectors are denoted by means of the transpose operator $(\cdot)^\top$. $\mathbf{1}$ (resp. $\mathbf{0}$) is a n -dimensional vector with each entry equals to one (resp. 0). e_i ($i = 1, \dots, n$) is the i th vector of the canonical basis of \mathbb{R}^n .

For any vector $x \in \mathbb{R}^n$, $\|x\|_1$ is its l_1 -norm, i.e. $\sum_{i=1}^n |x(i)|$. The l_1 -norm of any matrix M is defined by $\|M\|_1 := \max_i (\sum_{j=1}^n |M(i, j)|)$. Note that for any row-stochastic matrix M , $\|M\|_1 = 1$.

Let (Ω, \mathcal{F}) a measurable space. If μ is a finite signed measure on (Ω, \mathcal{F}) , the *total variation* $\text{Var}(\mu)$ of μ is defined by

$$\text{Var}(\mu) := \sup \left\{ \left| \int_{\Omega} f(\omega) d\mu(\omega) \right| : f \text{ is a } \mathcal{F}\text{-measurable function with } |f| \leq 1 \right\}.$$

The *total variation distance* $d_{\text{TV}}(\mu_1; \mu_2)$ between two finite signed measures μ_1, μ_2 on (Ω, \mathcal{F}) is

$$d_{\text{TV}}(\mu_1; \mu_2) := \text{Var}(\mu_1 - \mu_2).$$

Let \mathcal{B} be the Borel σ -algebra on \mathbb{R} and \mathcal{B}_+ its restriction to $[0, \infty[$. For $\omega \in \Omega$ fixed, any real-valued increasing function $\{A_t(\omega)\}_{t \geq 0}$ defines a positive measure $A(\omega, ds)$ on $(\mathbb{R}_+, \mathcal{B}_+)$. If $A_t(\omega) < +\infty$, then $A(\omega, ds)$ is finite on $[0, t]$. If $A^1(\omega), A^2(\omega)$ define two locally finite measures on $(\mathbb{R}_+, \mathcal{B}_+)$, then for any $t > 0$, $\text{Var}(A^1 - A^2)_t$ denotes the variation of $A^1(\omega, ds) - A^2(\omega, ds)$ on $([0, t], \mathcal{B}_+ \cap [0, t])$.

2 A Poisson approximation for a time-inhomogeneous MAP

2.1 A class of non-homogeneous Markovian Arrival Processes

In this part, we are concerned with the definition of a specific time-inhomogeneous MAP. We refer to [15, 10, 6] for basic properties of non-homogeneous Markov processes. We consider a càdlàg bivariate Markov process $(N, X) := \{(N_t, X_t)\}_{t \geq 0}$ with $(N_t, X_t) \in \mathbb{N} \times \mathcal{U}$ and \mathcal{U} is the finite set $\{e_i, i = 1, \dots, n\}$. The paths of $N := \{N_t\}_{t \geq 0}$ are assumed to be nondecreasing. The internal filtration of (N, X) is denoted by \mathbb{F} , i.e. $\mathbb{F} := \{\mathbb{F}_t\}_{t \geq 0}$ with $\mathbb{F}_t := \sigma(N_s, X_s, s \leq t)$. The

transition probabilities matrices $P(s, t)$ for (N, X) are defined by: for $e_i, e_j \in \mathcal{U}$, $m, k \in \mathbb{N}$ with $m \leq k$ and $0 \leq s \leq t$

$$P(s, t; (m, e_i), (k, e_j)) := \mathbb{P}\{N_t = k, X_t = e_j \mid N_s = m, X_s = e_i\}.$$

These transition probabilities matrices are supposed to be additive in the first component, that is, they satisfy for $m \leq k$ and $0 \leq s \leq t$

$$P(s, t; (m, e_i), (k, e_j)) = P(s, t; (0, e_i), (k - m, e_j)). \quad (2)$$

Thus, (N, X) is an instance of a Markov Additive Process [14] and has the following properties. $\{X_t\}_{t \geq 0}$ is a \mathbb{F} -Markov process with transition probabilities

$$P^*(s, t; (e_i, e_j)) = P(s, t; (0, e_i), (\mathbb{N}, e_j))$$

for every $e_i, e_j \in \mathcal{U}$ and $s \leq t$. N has conditionally independent increments given $\sigma(X_s, s \geq 0)$, i.e. the two σ -algebra $\sigma(N_t - N_s, t \geq s)$ and \mathbb{F}_s are conditionally independent given $\sigma(X_s, s \geq 0)$. Note that Breuer [2] has generalized analytical properties of homogeneous MAP stated in [14] to the class of non-homogeneous Markov additive processes which are Markovian jump processes according to [6].

Our main assumption concerns the existence of the so-called infinitesimal intensities $\mathcal{G}(t; \cdot, \cdot)$ for (N, X) . Intuitively, this means that

$$P(t, t + dt; z_1, z_2) = \begin{cases} \mathcal{G}(t; z_1, z_2)dt + o(dt) & \text{if } z_2 \neq z_1 \\ 1 + \mathcal{G}(t; z_1, z_1)dt + o(dt) & \text{if } z_2 = z_1. \end{cases} \quad (3)$$

The additivity property (2) of the transition probabilities implies that the infinitesimal intensities satisfy a similar condition: for $k \leq m$ and $t \geq 0$

$$\mathcal{G}(t; (m, e_i), (k, e_j)) = \mathcal{G}(t; (0, e_i), (k - m, e_j)). \quad (4)$$

Then, to define our specific Markov additive process, we will assume that, for $e_i, e_j \in \mathcal{U}$ and $t \geq 0$,

$$\mathcal{G}(t; (0, e_i), (k, e_j)) = 0 \quad \text{if } k > 1$$

so that N is the counting process of a simple point process. The precise statement is in the spirit of a martingale problem.

Definition 1 $(N, X) := \{(N_t, X_t)\}_{t \geq 0}$ with $(N_t, X_t) \in \mathbb{N} \times \mathcal{U}$ is said to be a non-homogeneous Markovian Arrival Process (NHMAP) if there exist matrices $D_k(t) := (D_k(t; i, j))_{i, j=1}^n$ $k = 0, 1$ such that :

1. for any i, j , $(1 - \delta_{j,i})D_0(t; i, j) \geq 0$, $D_1(t; i, j) \geq 0$ and $\sum_{j=1}^n (D_0(t; i, j) + D_1(t; i, j)) = 0$;
2. for any i, j , the functions $t \mapsto D_0(t; i, j)$, $t \mapsto D_1(t; i, j)$ are locally integrable;
3. for any $(k, e_i) \in \mathbb{N} \times \mathcal{U}$, the process

$$M_t((k, e_i)) := 1_{\{(N_t, X_t) = (k, e_i)\}} - \int_0^t \{1_{\{N_{s-} = k-1\}} X_{s-} D_1(s) + 1_{\{N_{s-} = k\}} X_{s-} D_0(s)\} e_i^\top ds$$

is a \mathbb{F} -martingale.

Remark 2 The special structure of Markov process (N, X) gives, for $i, j = 1, \dots, n$ and $k \geq 0$,

$$D_1(t; i, j) = \mathcal{G}(t; (k, e_i), (k+1, e_j)) \quad D_0(t; i, j) = \mathcal{G}(t; (k, e_i), (k, e_j)) \text{ if } i \neq j.$$

Therefore, the third condition in Definition 1 expresses the requirement for the process

$$M_t((k, e_i)) := 1_{\{(N_t, X_t) = (k, e_i)\}} - \int_0^t \mathcal{G}(s; (N_{s-}, X_{s-}), (k, e_i)) ds \quad (5)$$

to be a \mathbb{F} -martingale for every $(k, e_i) \in \mathbb{N} \times \mathcal{U}$. As in [10, 15], this condition is shown to be equivalent to the transition probabilities satisfy the Chapman-Kolmogorov equations.

Remark 3 Condition 2. in Definition 1 implies that $|D_0(t; i, j)| \leq \gamma_0(t)$, $D_1(t; i, j) \leq \gamma_1(t)$ where $\gamma_0(t)$ and $\gamma_1(t)$ are locally integrable. Then, Conditions 1-2 assert there exists a unique solution to the Chapman-Kolmogorov equations and we can construct a unique càdlàg Markov process having these infinitesimal intensities [5]. Note that, at time t fixed, the intensity matrix $\mathcal{G}(t)$ of (N, X) has the standard structure of the intensity matrix of a time-homogeneous MAP using the lexicographic ordering on the state space $\mathbb{N} \times \mathcal{U}$ (e.g. see [13]).

It follows from Definition 1 and (4) that $\{X_t\}_{t \geq 0}$ is a finite non-homogeneous Markov process with intensity matrices $\{Q^*(t)\}_{t \geq 0}$ given, for $t \geq 0$, by

$$Q^*(t) = D_0(t) + D_1(t).$$

We refer to [1, 11] for a treatment of the martingale approach to point processes. We only introduce the needed material for our purpose. Let $N_t(z_1, z_2)$ be the number of transitions of (N, X) from state z_1 to state z_2 on $[0, t]$. Since the \mathbb{F} -adapted process $\{1_{\{(N_{t-}, X_{t-}) = z_1\}}\}_{t \geq 0}$ is continuous to the left and bounded, the process $\mathcal{M}(z_1, z_2) := (\mathcal{M}_t(z_1, z_2))_{t \geq 0}$ defined by

$$\mathcal{M}_t(z_1, z_2) := \int_0^t 1_{\{(N_{s-}, X_{s-}) = z_1\}} dM_s(z_2).$$

is a \mathbb{F} -martingale for any $z_1, z_2 \in \mathbb{N} \times \mathcal{U}$. We obtain from (5) that the \mathbb{F} -martingale $\mathcal{M}(z_1, z_2)$ has the form $N(z_1, z_2) - A(z_1, z_2)$ where, for $t \geq 0$,

$$A_t(z_1, z_2) := \int_0^t 1_{\{(N_{s-}, X_{s-}) = z_1\}} \mathcal{G}(s; z_1, z_2) ds.$$

Using the specific form of the infinitesimal intensities of (N, X) , the non-zero $A(z_1, z_2)$ are

$$\begin{aligned} A_t((k, e_i), (k, e_j)) &= \int_0^t 1_{\{(N_{s-}, X_{s-}) = (k, e_i)\}} D_0(s; i, j) ds, \\ A_t((k, e_i), (k+1, e_j)) &= \int_0^t 1_{\{(N_{s-}, X_{s-}) = (k, e_i)\}} D_1(s; i, j) ds. \end{aligned}$$

Next, N is the counting process of transitions of (N, X) between pairs of states in the set $\mathcal{T} := \{((k, e_i); (k+1, e_j)), e_i, e_j \in \mathcal{U}, k \in \mathbb{N}\}$, i.e. $N_t = \sum_{(z_1, z_2) \in \mathcal{T}} N_t(z_1, z_2)$. Therefore, the \mathbb{F} -compensator of N is

$$A_t = \sum_{(z_1, z_2) \in \mathcal{T}} A_t(z_1, z_2) = \int_0^t X_{s-} D_1(s) \mathbf{1}^\top ds. \quad (6)$$

Note that N has a \mathbb{F} -intensity process $\{\lambda_t\}_{t \geq 0}$ given by

$$\lambda_t := X_{t-} D_1(t) \mathbf{1}^\top. \quad (7)$$

Let \mathbb{F}^N be the internal filtration of the counting process N . The \mathbb{F}^N -compensator \hat{A} and the \mathbb{F}^N -intensity $\{\hat{\lambda}_t\}_{t \geq 0}$ of N are from (6), (7) and [1, Chap 2, Th 14]

$$\hat{A}_t = \int_0^t \mathbb{E}[\lambda_s \mid \mathbb{F}_{s-}^N] ds = \int_0^t \hat{X}_{s-} D_1(s) \mathbf{1}^\top ds \quad \text{with} \quad \hat{\lambda}_t = \hat{X}_{t-} D_1(t) \mathbf{1}^\top \quad (8)$$

and $\hat{X}_t := \mathbb{E}[X_t \mid \mathbb{F}_t^N]$.

A Poisson process modulated by a non-homogeneous finite Markov process with intensity matrices $\{Q(t)\}_{t \geq 0}$ is a standard point process for various applications. This model is a special instance of an NHMAP with $D_0(t) := Q(t) - \text{diag}(\lambda(i))$ and $D_1(t) := \text{diag}(\lambda(i))$. The positive scalar $\lambda(i)$ is the intensity of the Poisson process when the underlying Markov process is in the state e_i . We retrieve from (7) the well-known expression $\sum_i \lambda(i) \mathbf{1}_{\{X_{t-}=e_i\}}$ of the \mathbb{F} -intensity for such a conditional Poisson process.

2.2 Model of perturbation and the convergence result

We introduce perturbed functions matrices $D_k^{(\varepsilon)}(t)$, $k = 0, 1$ on which the basic assumptions are as follows.

(AS1) The matrices $D_0^{(\varepsilon)}(t)$ and $D_1^{(\varepsilon)}(t)$ defined below, satisfy Definition 1

$$D_0^{(\varepsilon)}(t) = Q(t) + \varepsilon B_0(t) \quad \text{and} \quad D_1^{(\varepsilon)}(t) = \varepsilon B_1(t)$$

where, for every t , $Q(t)$ is a $n \times n$ -matrix of measurable functions with

$$(1 - \delta_{ij})Q(t; i, j) \geq 0, \quad Q(t) \mathbf{1}^\top = \mathbf{0}, \quad \sup_t \|Q(t)\|_1 < \infty$$

and $B_0(t), B_1(t)$ are matrices of measurable functions such that

$$(B_0(t) + B_1(t)) \mathbf{1}^\top = \mathbf{0}, \quad k = 0, 1 : \sup_t \|B_k(t)\|_1 < \infty.$$

Under (AS1), the family $\{Q(t)\}_{t \geq 0}$ properly defines a family of transition matrices from the Chapman-Kolmogorov equations. The NHMAP $(N^{(\varepsilon)}, X^{(\varepsilon)})$ is defined by the matrices $\{D_0^{(\varepsilon)}(t), D_1^{(\varepsilon)}(t), t \geq 0\}$ of (AS1). The Markov process $X^{(\varepsilon)} := \{X_t^{(\varepsilon)}\}_{t \geq 0}$ has intensity matrices $\{Q^{*,\varepsilon}(t)\}_{t \geq 0}$ where

$$Q^{*,\varepsilon}(t) := D_0^{(\varepsilon)}(t) + D_1^{(\varepsilon)}(t) = Q(t) + \varepsilon R(t) \quad (9)$$

with $R(t) \mathbf{1}^\top = \mathbf{0}$ and the family of matrices $\{R(t)\}_{t \geq 0}$ is uniformly bounded.

The second assumption is on the rate of convergence, as t tends to $+\infty$, of $Q(t)$ and $B_1(t)$ to some irreducible intensity matrix Q and matrix B_1 respectively.

(AS2) There exist matrices B_k ($k = 0, 1$), an irreducible intensity matrix Q with stationary distribution π , reals $\alpha, \beta > 1$ such that

$$\lim_{t \rightarrow +\infty} (2t)^\alpha \|Q(t) - Q\|_1 = 0, \quad \lim_{t \rightarrow +\infty} t^\beta \|B_1(t) - B_1\|_1 = 0 \quad \text{and} \quad \lim_{t \rightarrow +\infty} \|B_0(t) - B_0\|_1 = 0.$$

Under (AS2), the matrix $Q^{*,\varepsilon}(t)$ in (9) converges as t tends to $+\infty$.

From now, we consider the counting process $C^{(\varepsilon)} := \{C_t^{(\varepsilon)}\}_{t \geq 0}$ defined by

$$C_t^{(\varepsilon)} := N_{t/\varepsilon}^{(\varepsilon)} \quad t \geq 0. \quad (10)$$

For any $T > 0$, let $\mathbb{P}_{C^{(\varepsilon)},[0,T]}$ and $\mathbb{P}_{\mathcal{P},[0,T]}$ be the probability distributions of $C^{(\varepsilon)}$ and of a homogeneous Poisson process $\mathcal{P} := (\mathcal{P}_t)_{t \geq 0}$ on the space of all counting measures on $[0, T]$. In this section, we state that $\mathbb{P}_{C^{(\varepsilon)},[0,T]}$ converges to $\mathbb{P}_{\mathcal{P},[0,T]}$ at rate ε for the total variation distance. This is based on the following estimate of the distance between $\mathbb{P}_{C^{(\varepsilon)},[0,T]}$ and $\mathbb{P}_{\mathcal{P},[0,T]}$ [9, Th 2]

$$d_{\text{TV}}(\mathbb{P}_{C^{(\varepsilon)},[0,T]}; \mathbb{P}_{\mathcal{P},[0,T]}) \leq 2\mathbb{E} \text{Var}(\widehat{A}^{(\varepsilon)} - A)_T \quad (11)$$

where $\widehat{A}^{(\varepsilon)}$, A are the compensators of $C^{(\varepsilon)}$ and \mathcal{P} with respect to their internal filtration.

Theorem 4 *Assume (AS1-2). Let $C^{(\varepsilon)}$ be defined in (10) and $\mathcal{P} := (\mathcal{P}_t)_{t \geq 0}$ be a homogeneous Poisson process with intensity $\lambda := \pi B_1 \mathbf{1}^\top$. For any $T > 0$, there exists a constant κ_T such that*

$$d_{\text{TV}}(\mathbb{P}_{C^{(\varepsilon)},[0,T]}; \mathbb{P}_{\mathcal{P},[0,T]}) \leq \kappa_T \varepsilon.$$

Remark 5 *The order of the convergence rate in Theorem 4 cannot be improved in general. This follows from [4, Section 5. Ex 1], where the authors report a lower bound for the distance in variation that has order 1 in ε for a Poisson process modulated by a 2-states homogeneous Markov process.*

Note that κ_T is quadratic in T in general. The dependence is linear when we have $Q^{,\varepsilon}(t) = Q(t)$ for $t \geq 0$. This happens when the NHMAP is a Poisson process modulated by a Markov process for instance.*

3 Poisson-type approximation for a \mathcal{G} -marking of our NHMAP

In this section, we briefly discuss the case of marked arrivals. That is, we consider a sequence $\{(T_k, Z_k)\}_{k \geq 1}$ of random variables (assumed to be defined on the same filtered probability space) where: $\{T_k\}_{k \geq 1}$ ($T_0 := 0$) are the epochs of arrivals in a NHMAP and Z_k ($k \geq 1$) is the mark or value associated with the arrival time T_k . The archetype of such process is the standard (homogeneous) Batch Markovian Arrival Processes where Z_k are \mathbb{N} -valued (e.g. see [12]). A time-inhomogeneous batch Markovian arrival process may be defined from a natural generalization of Definition 1, a family of matrices $\{D_l(t)\}_{l \geq 0}$ replacing the two matrices $D_0(t), D_1(t)$. We omit the details.

Here, Z_k 's are assumed to be real-valued. We can associate to $\{(T_k, Z_k)\}_{k \geq 1}$ an integer random measure $\mu(\omega; \cdot, \cdot)$ on $(\mathbb{R}_+ \times \mathbb{R}, \mathcal{B}_+ \times \mathcal{B})$ defined by

$$\mu(\omega; dt, dx) = \sum_{k \geq 1} 1_{\{T_k(\omega) < \infty\}} \delta_{T_k(\omega), Z_k(\omega)}(dt, dx).$$

The internal filtration of μ is denoted by $\mathbb{F}^\mu := (\mathcal{F}_t^\mu)_{t \geq 0}$ with $\mathcal{F}_t^\mu := \sigma(\mu([0, s] \times B) : s \leq t, B \in \mathcal{B})$. Since $\mathbb{E}[N_t] = \mathbb{E}[\mu(\omega; [0, t] \times \mathbb{R})] < +\infty$, it is well-known that we can associate with μ an unique random measure ν such that: for any $B \in \mathcal{B}$ (1) $\{\nu([0, t] \times B)\}_{t \geq 0}$ is a \mathbb{F}^μ -predictable process and (2) $(\mu([0, t] \times B) - \nu([0, t] \times B))_{t \geq 0}$ is a \mathbb{F}^μ -martingale. ν is called

the \mathbb{F}^μ -compensator of μ or $\{(T_k, Z_k)\}_{k \geq 1}$. Here $\nu(\omega; \cdot, \cdot)$ is a locally finite positive measure on $(\mathbb{R}_+ \times \mathbb{R}, \mathcal{B}_+ \times \mathcal{B})$.

A family $\mathcal{G} := \{G(s, dx), s \geq 0\}$ of probability measures on \mathbb{R} such that $s \mapsto G(s, E)$ is \mathcal{B}_+ measurable for any $E \in \mathcal{B}$, is called a *stochastic kernel* from \mathbb{R}_+ into \mathbb{R} . We only consider a \mathcal{G} -marking μ of $\{T_k\}_{k \geq 1}$ as defined in [11, p189].

Definition 6 Let \mathcal{G} be a stochastic kernel from \mathbb{R}_+ into \mathbb{R} . $\{(T_k, Z_k)\}_{k \geq 1}$ or μ is said to be a \mathcal{G} -marking of the point process $\{T_k\}_{k \geq 1}$ if Z_1, Z_2, \dots are conditionally independent given $\{T_k\}_{k \geq 1}$ and

$$\mathbb{P}\{Z_k \in dx \mid \sigma(T_m, m \geq 1)\} = G(T_k, dx) \quad \mathbb{P}\text{-as on } \{T_k < +\infty\}.$$

If $G(t, dx) = G(dx)$, \mathcal{G} is said to be an independent marking of $\{T_k\}_{k \geq 1}$. We emphasize that a \mathcal{G} -marking μ is characterized by the following specific form of the so-called “disintegration property” of its \mathbb{F}^μ -compensator ν [11, Cor 6.2.6]:

$$\nu(dt, dx) = G(t, dx)A(dt) \quad (12)$$

where the process A is the \mathbb{F}^N -compensator of the underlying point process $\{T_k\}_{k \geq 1}$.

The probability measure $G(s, dx)$ is assumed to be absolutely continuous wrt the Lebesgue measure. If $\tilde{\mu}$ is the \mathcal{G} -marking of a homogeneous Poisson process with intensity λ , then the following simple inequality is easily obtained from (12): for $T > 0$,

$$\mathbb{E}\text{Var}(\nu - \tilde{\nu})_T \leq \mathbb{E}\text{Var}(A - \tilde{A})_T \quad (13)$$

where A is defined in (6), $\tilde{A} := (\lambda t)_{t \geq 0}$ and $\text{Var}(\nu - \tilde{\nu})_T$ is the variation of $\nu - \tilde{\nu}$ on $[0, T] \times \mathbb{R}$. Then, if we combine the result of [9, Th 2] with inequality (13), we obtain the following lemma.

Lemma 7 Fix $T > 0$. If $\mathbb{P}_{\mu, [0, T]}$ (resp. $\mathbb{P}_{\tilde{\mu}, [0, T]}$) is the probability distribution of the \mathcal{G} -marking μ (resp. $\tilde{\mu}$) of an NHMAP (resp. of a homogeneous Poisson process) in the space of all counting measures on $[0, T] \times \mathbb{R}$, then

$$d_{\text{TV}}(\mathbb{P}_{\mu, [0, T]}; \mathbb{P}_{\tilde{\mu}, [0, T]}) \leq 2\mathbb{E}\text{Var}(A - \tilde{A})_T.$$

Therefore, the convergence of the \mathcal{G} -marking of N to that of a Poisson process for the total variation distance, may be deduced from the convergence in total variation of the underlying point processes.

Remark 8 Inequality (13) also holds when $G(s, dx)$ is a probability distribution on a discrete set.

Let us consider the perturbed arrival process $N^{(\varepsilon)}$ which depends on the parameter ε as in Subsection 2.2. Let $\mu^{(\varepsilon)}$ be a \mathcal{G} -marking of $\{N_{t/\varepsilon}^{(\varepsilon)}\}_{t \geq 0}$ with a stochastic kernel \mathcal{G} which does not depend on ε . Under the conditions of Theorem 4, we obtain from Theorem 4 and Lemma 7 that, for every $T > 0$, there exists a constant κ_T such that

$$d_{\text{TV}}(\mathbb{P}_{\mu^{(\varepsilon)}, [0, T]}; \mathbb{P}_{\tilde{\mu}, [0, T]}) \leq \kappa_T \varepsilon \quad (14)$$

where $\tilde{\mu}$ is the \mathcal{G} -marking of the Poisson process with intensity $\pi B_1 \mathbf{1}^\top$.

Let us associate with $f(s, x) = x$,

$$U_t := \mu([0, t] \times \mathbb{R})(f) = \sum_{k \geq 1} Z_k 1_{\{T_k \leq t\}} \quad \text{and} \quad \tilde{U}_t := \tilde{\mu}([0, t] \times \mathbb{R})(f). \quad (15)$$

If $\tilde{\mu}$ is an independent marking of the Poisson process with intensity $\pi B_1 \mathbf{1}^\top$ then \tilde{U} is a Compound Poisson process. Since $N_t(\omega) < +\infty$, U_t has bounded variation on any compact set of \mathbb{R}_+ . And it follows from (14) that $\{U_t^{(\varepsilon)}\}_{t \geq 0}$ associated with $\mu^{(\varepsilon)}$ as in (15), converges in total variation to \tilde{U} over any $[0, T]$.

4 Proofs

Let us recall that $\hat{X}_t := \mathbb{E}[X_t \mid \mathbb{F}_t^N] = (\mathbb{P}\{X_t = e_i \mid \mathbb{F}_t^N\})_{i=1}^n$. The first step is to derive a filtering equation for $\{\hat{X}_t\}_{t \geq 0}$. Using the material in Subsection 2.1, the next lemma may be derived from [1, R7].

Lemma 9 *Let $X := \{X_t\}_{t \geq 0}$ be the second component of an NHMAP (N, X) . Let α be the probability distribution of X_0 . Then, for every $t \geq 0$,*

$$\hat{X}_t = \alpha + \int_0^t \hat{X}_s Q^*(s) ds + \int_0^t v_{s-} (dN_s - \hat{\lambda}_s ds)$$

where $\{\hat{\lambda}\}_{t \geq 0}$ is the \mathbb{F}^N -intensity of N given in (8) and for $s \geq 0$

$$v_{s-} := \frac{\hat{X}_{s-} D_1(s)}{\hat{\lambda}_s} - \hat{X}_{s-}.$$

Let $\{P^*(s, t), 0 \leq s \leq t\}$ be the transition matrices of the non-homogeneous Markov chain X . Then using a “variation of constants formula” [3, Chap 3], we obtain the following representation of \hat{X}_t from Lemma 9

$$\hat{X}_t = \alpha P^*(0, t) + \int_0^t v_{s-} P^*(s, t) (dN_s - \hat{\lambda}_s ds). \quad (16)$$

The proof of Theorem 4 is based on the two following lemmas. These lemmas are related to the perturbed NHMAP model $\{(N_t^{(\varepsilon)}, X_t^{(\varepsilon)})\}_{t \geq 0}$ defined from (AS1). The assumption (AS1) asserts, in particular, that $\{Q(t)\}_{t \geq 0}$ defines a family of transition matrices $\{P(s, t), 0 \leq s \leq t\}$. The first lemma gives a simple estimate of the l_1 -distance to $\mathbf{1}^\top \pi$ of the transition matrices of $X^{(\varepsilon)}$ in terms of the corresponding distance for the transition matrices $\{P(s, t), 0 \leq s \leq t\}$. The second lemma is on the control of the (expected) l_1 -distance of the filtered state vector $\widehat{X^{(\varepsilon)}}_t = \mathbb{E}[X_t^{(\varepsilon)} \mid \mathbb{F}_t^{N^{(\varepsilon)}}]$ to vector π . This distance naturally appears in evaluating the variations between the compensators of $\{N_{t/\varepsilon}^{(\varepsilon)}\}_{t \geq 0}$ and of the Poisson process with intensity $\pi B_1 \mathbf{1}^\top$. Then, the control of these variations allows the use of inequality (11) for proving Theorem 4.

Lemma 10 *Assume (AS1-2). Let $\{P(s, t), 0 \leq s \leq t\}$ and $\{P^{*,\varepsilon}(s, t), 0 \leq s \leq t\}$ be the transition matrices defined from $\{Q(t)\}_{t \geq 0}$ and $\{Q^{*,\varepsilon}(t)\}_{t \geq 0}$ respectively. Then, for $0 \leq s \leq t$,*

$$\|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 \leq \|P(s, t) - \mathbf{1}^\top \pi\|_1 + \kappa \varepsilon \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr$$

where κ does not depend on ε and s, t .

Proof The triangle inequality allows us to write

$$\|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 \leq \|P^{*,\varepsilon}(s, t) - P(s, t)\|_1 + \|P(s, t) - \mathbf{1}^\top \pi\|_1.$$

It remains to estimate the norm $\|P^{*,\varepsilon}(s, t) - P(s, t)\|_1$. Recall that $Q^{*,\varepsilon}(t) = Q(t) + \varepsilon R(t)$ with $R(t)\mathbf{1}^\top = \mathbf{0}$ (see (9)). First, set for $0 \leq s \leq t$

$$K(s, t) := P^{*,\varepsilon}(s, t) - P(s, t).$$

Note that $K(s, s) = \mathbf{0}$. We find from the forward Chapman-Kolmogorov equation that, for $0 \leq s < t$,

$$\begin{aligned} K(s, t) &= \int_s^t (P^{*,\varepsilon}(s, r)Q^{*,\varepsilon}(r) - P(s, r)Q(r))dr \\ &= \int_s^t P^{*,\varepsilon}(s, r)(Q^{*,\varepsilon}(r) - Q(r))dr + \int_s^t K(s, r)Q(r)dr \\ &= \int_s^t P^{*,\varepsilon}(s, r)\varepsilon R(r)dr + \int_s^t K(s, r)Q(r)dr. \end{aligned}$$

Then, using a “variation of constants formula”, we obtain for $0 \leq s \leq t$

$$K(s, t) = \varepsilon \int_s^t P^{*,\varepsilon}(s, r)R(r)P(r, t)dr.$$

Since $R(r)\mathbf{1}^\top = \mathbf{0}$ for $r \geq 0$, we can write for $0 \leq s \leq t$

$$K(s, t) = \varepsilon \int_s^t P^{*,\varepsilon}(s, r)R(r)(P(r, t) - \mathbf{1}^\top \pi)dr.$$

Since $\|P^{*,\varepsilon}(s, r)\|_1 = 1$ and $\|R(r)\|_1$ is uniformly bounded from (AS1), we deduce from the equality above that, there exists a constant κ which does not depend on ε, s, t such that, for $0 \leq s \leq t$,

$$\|P^{*,\varepsilon}(s, t) - P(s, t)\|_1 = \|K(s, t)\|_1 \leq \varepsilon \kappa \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr.$$

Lemma 11 Assume (AS1). Let $X^{(\varepsilon)} := \{X_t^{(\varepsilon)}\}_{t \geq 0}$ be the second component of the perturbed NHMAP $(N^{(\varepsilon)}, X^{(\varepsilon)})$ and $\widehat{X}^{(\varepsilon)}_t := \mathbb{E}[X_t^{(\varepsilon)} \mid \mathbb{F}_t^{N^{(\varepsilon)}}]$. The following inequality holds for any $t \geq 0$,

$$\mathbb{E}\|\widehat{X}^{(\varepsilon)}_t - \pi\|_1 \leq \|P(0, t) - \mathbf{1}^\top \pi\|_1 + \kappa \varepsilon \int_0^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr + \kappa \varepsilon^2 \int_0^t \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr ds \quad (17)$$

where κ is a constant which does not depend on ε, t and where $P(s, t)$'s are the transition matrices associated with the family of intensity matrices $\{Q(t)\}_{t \geq 0}$.

Proof. Let α be the probability distribution of X_0 or $X_0^{(\varepsilon)}$. Applying Lemma 9 to the perturbed NHMAP $(N^{(\varepsilon)}, X^{(\varepsilon)})$, we deduce that $\{\widehat{X}^{(\varepsilon)}_t\}_{t \geq 0}$ satisfies an equation of the form (16)

$$\widehat{X}^{(\varepsilon)}_t = \alpha P^{*,\varepsilon}(0, t) + \int_0^t v_{s-}^{(\varepsilon)} P^{*,\varepsilon}(s, t)(dN_s^{(\varepsilon)} - \widehat{\lambda}_s^{(\varepsilon)} ds) \quad (18)$$

where $\{P^{*,\varepsilon}(s, t), 0 \leq s \leq t\}$ are the transition matrices of $X^{(\varepsilon)}$ and

$$v_{s-}^{(\varepsilon)} = \frac{\widehat{X^{(\varepsilon)}}_{s-} D_1^{(\varepsilon)}(s)}{\widehat{\lambda_s^{(\varepsilon)}}} - \widehat{X^{(\varepsilon)}}_{s-} \quad \widehat{\lambda_s^{(\varepsilon)}} = \widehat{X^{(\varepsilon)}}_{s-} D_1^{(\varepsilon)}(s) \mathbf{1}^\top. \quad (19)$$

When $t = 0$, inequality (17) follows from $\widehat{X^{(\varepsilon)}}_0 = \alpha$, $\|\alpha\|_1 = 1$, $P^{*,\varepsilon}(0, 0) = P(0, 0) = I$ and

$$\|\alpha - \pi\|_1 = \|\alpha(P^{*,\varepsilon}(0, 0) - \mathbf{1}^\top \pi)\|_1 \leq \|\alpha\|_1 \|P^{*,\varepsilon}(0, 0) - \mathbf{1}^\top \pi\|_1 = \|P(0, 0) - \mathbf{1}^\top \pi\|_1.$$

Because $v_{s-}^{(\varepsilon)} \mathbf{1}^\top = 0$ for $s \geq 0$, and $\widehat{X^{(\varepsilon)}}_0 \mathbf{1}^\top = \alpha \mathbf{1}^\top = 1$, we can write from (18) that for $t > 0$

$$\widehat{X^{(\varepsilon)}}_t - \pi = \alpha(P^{*,\varepsilon}(0, t) - \mathbf{1}^\top \pi) + \int_0^t v_{s-}^{(\varepsilon)}(P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi)(dN_s^{(\varepsilon)} - \widehat{\lambda_s^{(\varepsilon)}} ds).$$

Since $\|v_{t-}^{(\varepsilon)}\|_1 \leq 2$ as the l_1 -norm of the difference of two stochastic vectors and $\|\alpha\|_1 = 1$, it follows that

$$\|\widehat{X^{(\varepsilon)}}_t - \pi\|_1 \leq \|P^{*,\varepsilon}(0, t) - \mathbf{1}^\top \pi\|_1 + 2 \int_0^t \|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 (dN_s^{(\varepsilon)} + \widehat{\lambda_s^{(\varepsilon)}} ds). \quad (20)$$

Since $\{\widehat{\lambda_t^{(\varepsilon)}}\}_{t \geq 0}$ is the $\mathbb{F}^{N^{(\varepsilon)}}$ -intensity of the counting process $N^{(\varepsilon)}$, we know that [1]

$$\mathbb{E} \int_0^t \|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 dN_s^{(\varepsilon)} = \mathbb{E} \int_0^t \|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 \widehat{\lambda_s^{(\varepsilon)}} ds.$$

Then, taking the expectation on both side of (20), we find for every $t > 0$

$$\mathbb{E} \|\widehat{X^{(\varepsilon)}}_t - \pi\|_1 \leq \|P^{*,\varepsilon}(0, t) - \mathbf{1}^\top \pi\|_1 + 4 \int_0^t \|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 \widehat{\lambda_s^{(\varepsilon)}} ds.$$

Hereafter, $\kappa > 0$ stands for a generic positive constant which does not depend on ε and t . Since $\|\widehat{X^{(\varepsilon)}}_{t-}\|_1 = 1$, it easily follows from (AS1) that, for $t \geq 0$,

$$\widehat{\lambda_t^{(\varepsilon)}} = \widehat{X^{(\varepsilon)}}_{t-} D_1^{(\varepsilon)}(t) \mathbf{1}^\top \leq \kappa \varepsilon.$$

Then, we obtain for $t > 0$

$$\mathbb{E} \|\widehat{X^{(\varepsilon)}}_t - \pi\|_1 \leq \|P^{*,\varepsilon}(0, t) - \mathbf{1}^\top \pi\|_1 + \kappa \varepsilon \int_0^t \|P^{*,\varepsilon}(s, t) - \mathbf{1}^\top \pi\|_1 ds.$$

Under (AS1), we have $Q^{*,\varepsilon}(t) = Q(t) + \varepsilon R(t)$ with $R(t) \mathbf{1}^\top = \mathbf{0}$ and Lemma 10 applies

$$\begin{aligned} \mathbb{E} \|\widehat{X^{(\varepsilon)}}_t - \pi\|_1 &\leq \|P(0, t) - \mathbf{1}^\top \pi\|_1 + \kappa \varepsilon \int_0^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr \\ &\quad + \kappa \varepsilon \int_0^t (\|P(s, t) - \mathbf{1}^\top \pi\|_1 + \kappa \varepsilon \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr) ds. \end{aligned}$$

Inequality (17) easily follows.

Proof of Theorem 4. The compensator $A := \{A_t\}_{t \geq 0}$ of the Poisson process $\{\mathcal{P}_t\}_{t \geq 0}$ with intensity $\pi B_1 \mathbf{1}^\top$, is $A_t = t\pi B_1 \mathbf{1}^\top$. Since the compensators $\widehat{A}^{(\varepsilon)}$ of $C^{(\varepsilon)}$ and A are absolutely continuous w.r.t. the Lebesgue measure, we find that

$$\text{Var}(\widehat{A}^{(\varepsilon)} - A)_t = \int_0^{t/\varepsilon} |\widehat{\lambda}_s^{(\varepsilon)} - \varepsilon \pi B_1 \mathbf{1}^\top| ds \quad (21)$$

where $\widehat{\lambda}_s^{(\varepsilon)}$ is defined in (19). Under (AS1), we obtain from (11), (21), (19) and the triangle inequality that

$$d_{\text{TV}}(\mathbb{P}_{C^{(\varepsilon)}, [0, T]}; \mathbb{P}_{\mathcal{P}, [0, T]}) \leq 2\varepsilon \mathbb{E} \int_0^{T/\varepsilon} |(\widehat{X}^{(\varepsilon)}_{s-} - \pi) B_1 \mathbf{1}^\top| ds + 2\varepsilon \mathbb{E} \int_0^{T/\varepsilon} |\widehat{X}^{(\varepsilon)}_{s-} (B_1(s) - B_1) \mathbf{1}^\top| ds. \quad (22)$$

Since B_1 and $\widehat{X}^{(\varepsilon)}$ are bounded, we have for some constant κ which does not depend on ε, s

$$|(\widehat{X}^{(\varepsilon)}_{s-} - \pi) B_1 \mathbf{1}^\top| \leq \kappa \|\widehat{X}^{(\varepsilon)}_{s-} - \pi\|_1 \quad \text{and} \quad |\widehat{X}^{(\varepsilon)}_{s-} (B_1(s) - B_1) \mathbf{1}^\top| \leq \kappa \|B_1(s) - B_1\|_1.$$

Under (AS2), we obtain that the second term in the right hand side member of (22) is bounded from above by $\kappa \varepsilon \int_0^{+\infty} \|B_1(s) - B_1\|_1 ds$. Thus, we obtain that, for some constant κ_1 which does not depend on ε, T

$$d_{\text{TV}}(\mathbb{P}_{C^{(\varepsilon)}, [0, T]}; \mathbb{P}_{\mathcal{P}, [0, T]}) \leq \kappa_1 \varepsilon \left(\mathbb{E} \int_0^{T/\varepsilon} \|\widehat{X}^{(\varepsilon)}_{s-} - \pi\|_1 ds + 1 \right).$$

Note that $\int_0^{T/\varepsilon} \|\widehat{X}^{(\varepsilon)}_{s-} - \pi\|_1 ds = \int_0^{T/\varepsilon} \|\widehat{X}^{(\varepsilon)}_s - \pi\|_1 ds$ since the paths of $\widehat{X}^{(\varepsilon)}$ have at most a finite number of discontinuities on any finite interval. From Fubini's theorem and Lemma 11, the expectation in the inequality above, may be bounded as follows: there exists a constant κ_2 that does not depend on ε, T such that, for $T > 0$,

$$\begin{aligned} \int_0^{T/\varepsilon} \mathbb{E} \|\widehat{X}^{(\varepsilon)}_t - \pi\|_1 dt &\leq \int_0^{T/\varepsilon} \|P(0, t) - \mathbf{1}^\top \pi\|_1 dt + \kappa_2 \varepsilon \int_0^{T/\varepsilon} \int_0^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr dt \\ &\quad + \kappa_2 \varepsilon^2 \int_0^{T/\varepsilon} \int_0^t \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr ds dt. \end{aligned}$$

Under (AS2), we deduce from (A.2), (A.3) and (A.4) that

$$\int_0^{T/\varepsilon} \mathbb{E} \|\widehat{X}^{(\varepsilon)}_t - \pi\|_1 dt \leq \kappa_{2, T}.$$

for some constant $\kappa_{2, T}$. The proof is complete.

A Rates of convergence for non-homogeneous Markov processes

A finite non-homogeneous Markov process $\{X_t\}_{t \geq 0}$ with transition matrices $\{P(s, t), 0 \leq s \leq t\}$ is said to be strongly ergodic if there exists a constant stochastic vector π such that

$$\forall s \geq 0, \quad \lim_{t \rightarrow +\infty} \|P(s, t) - \mathbf{1}^\top \pi\|_1 = 0.$$

If the previous limit is uniform in s , then the Markov process is said to be *uniformly strongly ergodic*. We have the following basic result.

Lemma 12 ([8]) *Let $\{X_t\}_{t \geq 0}$ be a finite non-homogeneous Markov process with intensity matrices $\{Q(t)\}_{t \geq 0}$ satisfying $\sup_t \|Q(t)\|_1 < +\infty$. Let Q be an irreducible intensity matrix with stationary distribution π . If $Q(t)$ converges to Q in norm l_1 as $t \rightarrow +\infty$, then $\{X_t\}_{t \geq 0}$ is uniformly strongly ergodic.*

Any homogeneous Markov process with irreducible intensity matrix Q is (strongly) ergodic. Then, the transition matrix $P(t) = \exp(Qt)$ has an exponential rate of convergence to $\mathbf{1}^\top \pi$ as $t \rightarrow +\infty$. But an exponential rate of convergence is not guaranteed in the non-homogeneous case. Indeed, the rate at which $P(s, t)$ converges to $\mathbf{1}^\top \pi$ as $t \rightarrow +\infty$ is governed by the rate at which $Q(t)$ converges to Q .

Lemma 13 ([8, Th 8]) *Under the conditions of Lemma 12, if $Q(t)$ converges to Q in norm l_1 with*

$$\lim_{t \rightarrow +\infty} (2t)^\alpha \|Q(t) - Q\|_1 = 0$$

for some $\alpha > 0$, then

$$\lim_{t \rightarrow +\infty} t^\alpha \sup_{s \geq 0} \|P(s, s+t) - \mathbf{1}^\top \pi\|_1 = 0.$$

If α is greater than 1 in the previous lemma, then the following integral converges

$$\sup_{s \geq 0} \int_0^{+\infty} \|P(s, s+t) - \mathbf{1}^\top \pi\|_1 dt \leq \int_0^{+\infty} \sup_{s \geq 0} \|P(s, s+t) - \mathbf{1}^\top \pi\|_1 dt \leq \kappa < +\infty \quad (\text{A.1})$$

The introduction of our assumption (AS2) is motivated by Lemma 13.

The following properties are easily deduced from (A.1). They are used in the proof of Theorem 4.

$$\forall r \geq 0, \quad \int_r^{+\infty} \|P(r, t) - \mathbf{1}^\top \pi\|_1 dt \leq \kappa \quad (\text{A.2})$$

$$\int_0^{T/\varepsilon} \int_0^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr dt \leq \frac{T}{\varepsilon} \kappa \quad (\text{A.3})$$

$$\int_0^{T/\varepsilon} \int_0^t \int_s^t \|P(r, t) - \mathbf{1}^\top \pi\|_1 dr ds dt \leq \left(\frac{T}{\varepsilon}\right)^2 \kappa. \quad (\text{A.4})$$

References

- [1] Bremaud, P. *Point Processes and Queues*. Springer Series in Statistics. Springer, 1981.
- [2] Breuer, L. On Markov-additive jump processes. *Queueing Systems*, 40:75–91, 2002.
- [3] Coddington, E.A. and Levinson, N. *Theory of Ordinary Differential Equations*. McGraw-Hill, 1955.

- [4] Di Masi, G.B. and Kabanov, Y.M. The strong convergence of two-scale stochastic systems and singular perturbations of filtering equations. *J. Math. Syst. Estim. Cont.*, 3:207–224, 1993.
- [5] Ethier, S.N. and Kurtz, T.G. *Markov Processes, characterization and convergence*. John Wiley & Sons, 1986.
- [6] Gikhman, I. and Skorokhod, A. V. *Introduction to the Theory of Random Processes*. Saunders, 1969.
- [7] Gravereaux, J.-B. and Ledoux, J. Poisson approximation for some point processes in reliability. *Adv. in Appl. Probab.*, 36:455–470, 2004.
- [8] Johnson, J. and Isaacson, D. Conditions for strong ergodicity using intensity matrices. *J. Appl. Probab.*, 25:34–42, 1988.
- [9] Kabanov, Yu.M. and Liptser, R.Sh. On convergence in variation of the distributions of multivariate point processes. *Z. Wahrsch. Verw. Gebiete*, 63:475–485, 1983.
- [10] Krylov, N.V. and Zatezalo, A. Filtering of finite-state time-nonhomogeneous Markov processes, a direct approach. *Appl. Math. Optim.*, 42:229–258, 2000.
- [11] Last, G. and Brandt, A. *Marked Point Processes on the Real Line*. Springer, 1995.
- [12] Lucantoni, D.M. The BMAP/G/1 queue: a tutorial. In L. Donatiello and R. Nelson, editors, *Performance Evaluation of Computer and Communication Systems*, LNCS No 729, pages 330–358. Springer-Verlag, 1993.
- [13] Neuts, M. F. *Structured Stochastic Matrices of M/G/1 Type and their Applications*. Marcel Dekker Inc., New-York and Basel, 1989.
- [14] Pacheco, A. and Prabhu, N.U. Markov-additive processes of arrivals. In J.H. Dshalalow, editor, *Adv. in Queueing Theory, Methods, and Open Problems*, pages 167–194. CRC Press, 1995.
- [15] Yin, G.G and Zhang, Q. *Continuous-time Markov chains and Applications*. Springer, 1998.